



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2021

---

## **Design of SNP markers for Aldabra giant tortoises using low coverage ddRAD-seq**

Çilingir, F G ; Hansen, Dennis ; Ozgul, Arpat ; Grossen, Christine

**Abstract:** The Aldabra giant tortoise (*Aldabrachelys gigantea*) is one of only two remaining giant tortoise species worldwide. Captive-bred *A. gigantea* are being used in rewilding projects in the Western Indian Ocean to functionally replace the extinct endemic giant tortoise species and restore degraded island ecosystems. Furthermore, large-scale translocations may become necessary as rising sea levels threaten the only wild population on the low-lying Aldabra Atoll. Critical management decisions would be greatly facilitated by insights on the genetic structure of breeding populations. We used a double-digest restriction-associated DNA sequencing (ddRAD-seq) approach to identify single nucleotide polymorphisms (SNP) among the wild population and two additional captive populations of *A. gigantea*. A total of 1674 unlinked, putatively neutral genome-wide SNPs were identified. The values of expected heterozygosity ranged from 0.33 to 0.5, whereas the minor allele frequency ranged from 0.20 to 0.5. These novel SNP markers will serve as useful tools for informing the conservation of *A. gigantea*.

DOI: <https://doi.org/10.1007/s12686-021-01225-4>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-206217>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Çilingir, F G; Hansen, Dennis; Ozgul, Arpat; Grossen, Christine (2021). Design of SNP markers for Aldabra giant tortoises using low coverage ddRAD-seq. *Conservation Genetics Resources*, 13(4):409-412. DOI: <https://doi.org/10.1007/s12686-021-01225-4>



# Design of SNP markers for Aldabra giant tortoises using low coverage ddRAD-seq

F. G. Çilingir<sup>1</sup> · D. Hansen<sup>1,2</sup> · A. Ozgul<sup>1</sup> · C. Grossen<sup>1</sup>

Received: 5 May 2021 / Accepted: 14 July 2021  
© The Author(s) 2021

## Abstract

The Aldabra giant tortoise (*Aldabrachelys gigantea*) is one of only two remaining giant tortoise species worldwide. Captive-bred *A. gigantea* are being used in rewilding projects in the Western Indian Ocean to functionally replace the extinct endemic giant tortoise species and restore degraded island ecosystems. Furthermore, large-scale translocations may become necessary as rising sea levels threaten the only wild population on the low-lying Aldabra Atoll. Critical management decisions would be greatly facilitated by insights on the genetic structure of breeding populations. We used a double-digest restriction-associated DNA sequencing (ddRAD-seq) approach to identify single nucleotide polymorphisms (SNP) among the wild population and two additional captive populations of *A. gigantea*. A total of 1674 unlinked, putatively neutral genome-wide SNPs were identified. The values of expected heterozygosity ranged from 0.33 to 0.5, whereas the minor allele frequency ranged from 0.20 to 0.5. These novel SNP markers will serve as useful tools for informing the conservation of *A. gigantea*.

**Keywords** ddRAD-seq · *Aldabrachelys gigantea* · SNPs · Giant tortoises

The Aldabra giant tortoise (*Aldabrachelys gigantea*) is one of only two extant giant tortoise species together with the Galapagos giant tortoise complex (*Chelonoidis niger*) (Turtle Taxonomy Working Group 2017). The natural distribution of *A. gigantea* is restricted to Aldabra Atoll, Seychelles, northeast of Madagascar. The species is the largest vertebrate in the ecosystem and fulfills prominent functional roles in shaping and sustaining large-scale landscape dynamics, such as herbivory, seed dispersal, and nutrient cycling (Hansen 2015; Hnatiuk et al. 1976; Merton et al. 1976). Accordingly, for the rewilding of several Western Indian Ocean Islands, captive-bred *A. gigantea* have been introduced to degraded ecosystems as non-native species to functionally replace extinct endemic species of giant tortoises (Griffiths et al. 2010, 2013; Hansen et al. 2010). Recent studies on feeding behavior, fecal analyses, and vegetation monitoring have shown that these replacement tortoises seem to be doing a good job in resurrecting herbivory

and seed dispersal interactions (Griffiths et al. 2010, 2013; Hansen et al. 2010). Depending on the effectiveness of these rewilding projects, spatially larger projects in Madagascar are planned, with pilot releases underway (Pedrono et al. 2013, 2020). Large-scale translocations may also become necessary for the protection of the species, as rising sea levels threaten the large, wild population on the low-lying Aldabra Atoll (Falcón et al. 2018).

Despite the increasing trend of application, a proper evaluation of project outcomes is still lacking from the majority of the rewilding programs (Nogués-Bravo et al. 2016; Svenning et al. 2016). Although some evaluations have looked at the functional roles of the introduced populations, none have assessed the genetic variation within both wild and rewilded populations. Incorporating genetic monitoring in the evaluation of rewilding projects is critical to enable the long-term survival of translocated populations (Stronen et al. 2019). Furthermore, the mating system of *A. gigantea* remains largely unknown, but such knowledge would highly facilitate the efficient management of captive populations.

To address these needs, we performed double digest restriction sequencing (ddRad-seq; Peterson et al. [2012]). We extracted DNA from 35 blood samples, representing all main islands of Aldabra Atoll, where wild *A. gigantea* occur (Malabar, Grande Terre, and Picard, N = 33), and two

✉ F. G. Çilingir  
goezde.cilingir@uzh.ch

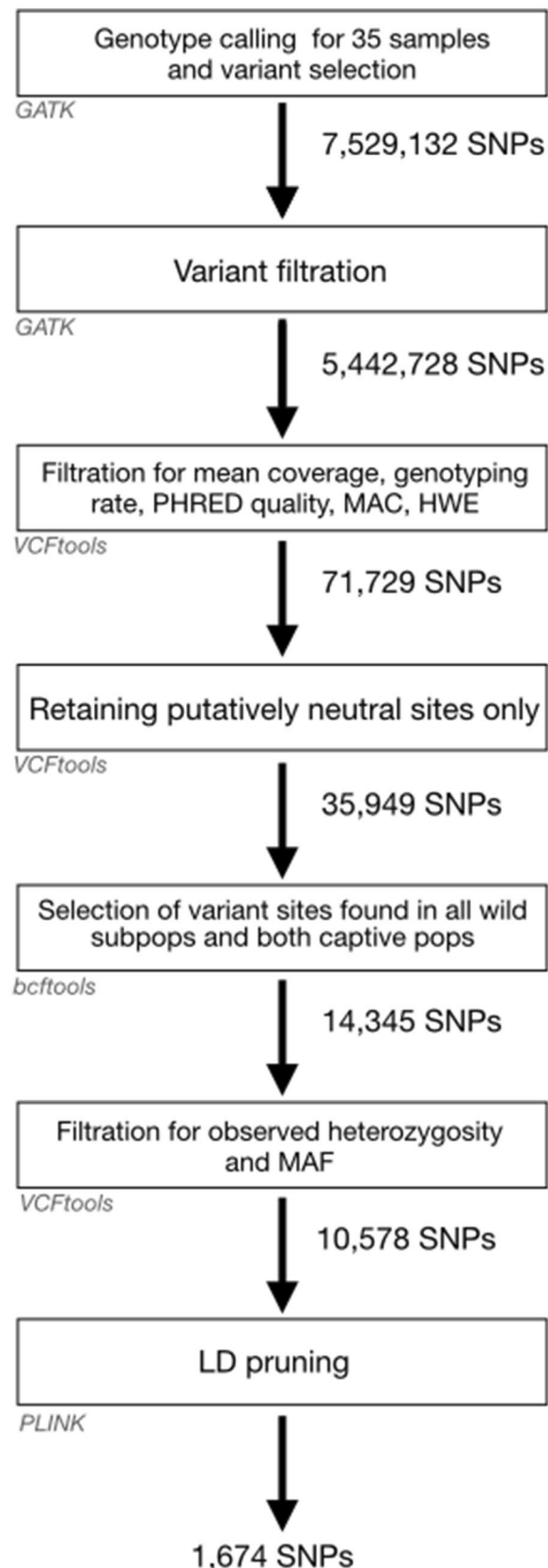
<sup>1</sup> Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland

<sup>2</sup> Zoological Museum, University of Zurich, 8006 Zurich, Switzerland

captive populations in Mauritius and Rodrigues (N=2, Supplementary Information 1&2). DNA extractions were performed using a beadex™ kit following the manufacturer's protocol. A single ddRad-seq library was prepared with the restriction enzymes EcoRI and BfaI and sequenced on an Illumina Miseq platform with a 300 bp paired-end run. A total of 24,942,560 raw reads were generated for the 35 individuals of which 94.3% were retained after demultiplexing with Stacks v2.53 (Rochette et al. 2019) and quality filtering with Trimmomatic v0.39 (Bolger et al. 2014).

Although there is currently no reference genome available for *A. gigantea*, we chose a reference-based over a de novo SNP discovery. Our decision was mainly based on the high mapping rate of 96.5% (range: 94.7–97.0%) attained when aligning the quality-filtered data to the *Gopherus evgoodei* genome (NCBI Accession PRJNA559383, the phylogenetically closest species that has a chromosome-level assembled genome available [Kehlmaier et al. 2019], using default settings of BWA-MEM version 0.7.17 [Li and Durbin 2009]). Furthermore, the use of an annotated genome made it possible to exclude coding regions in order to focus on putatively neutral SNPs (see below). Sequencing depth per sample was 2.28x (range 0.2–6.1x) again making a reference-based SNP discovery (here with the HaplotypeCaller module of GATK v4.1.7 [McKenna et al. 2010] in GVCF mode) more suitable than a de novo SNP discovery. Genotypes of all 35 samples were called jointly, using GATK GenotypeGVCFs. SNPs were preliminarily filtered using the GATK VariantFiltration function by following the GATK Best Practices recommendations (Van der Auwera et al. 2013). Eventually, 5,442,728 SNPs were filtered with VCFtools 0.1.17 (Danecek et al. 2011) to only retain sites with coverage between 2 and 10, and a genotyping rate > 50%. This resulted in a minimum coverage per called SNP of 48x. Sites with PHRED quality < 30, minor allele count (MAC) < 3, and deviating from Hardy Weinberg Equilibrium (HWE) with a p-value < 0.00001 were excluded. Next, SNPs located in the non-coding regions of the reference genome (hence putatively neutral) and polymorphic in all three wild subpopulations, as well as both of the captive populations were selected as candidate SNP markers. To avoid putative paralogs, SNPs with observed heterozygosity > 0.5 were removed, and SNPs with minor allele frequency (MAF) < 0.2, and > 0.8 were filtered out to maximize the potential information content. LD pruning was performed with the default settings of PLINK v1.9 (Purcell et al. 2007) except for a window size of 4 Gbp (to include all possible pairwise LD estimates). Finally, we obtained a total of 1674 genome-wide and putatively neutral SNPs (Fig. 1). (See also Supplementary Information 3 for all the details of the bioinformatic analyses).

The values of expected heterozygosity and the MAF ranged from 0.33 to 0.50 and from 0.20 to 0.50,



**Fig. 1** A flow diagram summarizing all the filtering steps that eventually yielded the total number of SNPs identified

respectively. The sequences spanning  $\pm 300$  bp of the SNP positions and the rest of the details about the 1674 genome-wide SNPs discovered for *A. gigantea* are summarized in Supplementary Information 4 and 5, respectively. These novel SNP markers will serve as a tool for the population structure analysis of *A. gigantea* and facilitate the management decisions regarding future rewilding plans using *A. gigantea* as important ecosystem engineers.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12686-021-01225-4>.

**Acknowledgements** We would like to thank the Seychelles Islands Foundation (SIF) fieldwork team and Richard Baxter for collecting the blood samples in Aldabra Atoll, and the SIF for their assistance in getting all the permits required to get the blood samples. We are also thankful to Owen Griffiths for his assistance in getting blood samples in Mauritius and Rodrigues, and to Erik Postma and Xenia Wietlisbach for providing the blood samples. Also, we would like to thank Claudia Michel for her contributions to the optimization of the ddRAD-seq protocol. All the wet-lab work and the sequencing efforts required for this study were conducted in the Genetic Diversity Center, ETH Zurich. F. Gözde Çilingir was funded by the Swiss Government Excellence Scholarship for Postdocs and the UZH Alumni Postdoctoral Fellowship during this study. Georges and Antoine Claraz Foundation contributed to the sequencing costs of this project.

**Funding** Open Access funding provided by Universität Zürich. F. Gözde Çilingir was funded by the Swiss Government Excellence Scholarship for Postdocs and the UZH Alumni Postdoctoral Fellowship during this study. Georges and Antoine Claraz Foundation contributed to the sequencing costs of this project.

**Data availability** Raw data will be provided upon request.

**Code availability** All bioinformatic codes used in this study are available in the Supplementary Information.

## Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158
- Falcón W, Baxter RP, Furrer S, Bauert M, Hatt JM, Schaepman-Strub G, Özgül A, Bunbury N, Claus M, Hansen DM (2018) Patterns of activity and body temperature of Aldabra giant tortoises in relation to environmental temperature. *Ecol Evol* 8(4):2108–2121
- Griffiths CJ, Jones CG, Hansen DM, Puttoo M, Tatayah RV, Müller CB, Harris S (2010) The use of extant non-indigenous tortoises as a restoration tool to replace extinct ecosystem engineers. *Restor Ecol* 18(1):1–7
- Griffiths CJ, Zuñel N, Jones CG, Ahamud Z, Harris S (2013) Assessing the potential to restore historic grazing ecosystems with tortoise ecological replacements. *Conserv Biol* 27(4):690–700
- Hansen DM (2015) Non-native megaherbivores: the case for novel function to manage plant invasions on islands. *AoB Plants* 7:plv085
- Hansen DM, Josh Donlan C, Griffiths CJ, Campbell KJ (2010) Ecological history and latent conservation potential: large and giant tortoises as a model for taxon substitutions. *Ecography* 33:272–284
- Hnatiuk RJ, Woodell SRJ, Bourn DM (1976) Giant tortoise and vegetation interactions on Aldabra atoll—Part 2: Coastal. *Biol Conserv* 9(4):305–316
- Kehlmaier C, Graciá E, Campbell PD, Hofmeyr MD, Schweiger S, Martínez-Silvestre A, Joyce W, Fritz U (2019) Ancient mitogenomics clarifies radiation of extinct Mascarene giant tortoises (*Cylindraspis* spp.). *Sci Rep* 9(1):17487
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303
- Merton LFH, Bourn DM, Hnatiuk RJ (1976) Giant tortoise and vegetation interactions on Aldabra atoll—Part 1: Inland. *Biol Conserv* 9(4):293–304
- Nogués-Bravo D, Simberloff D, Rahbek C, Sanders NJ (2016) Rewilding is the new Pandora's box in conservation. *Curr Biol* 26(3):R87–R91
- Pedrono M, Rambeloson E, Clausen A (2020) Giant tortoises make a comeback in Madagascar. *Nature* 587(7835):548
- Pedrono M, Griffiths OL, Clausen A, Smith LL, Griffiths CJ, Wilmé L, Burney DA (2013) Using a surviving lineage of Madagascar's vanished megafauna for ecological restoration. *Biol Conserv* 59:501–506
- Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7(5):e37135
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575

- Rochette NC, Rivera-Colón AG, Catchen JM (2019) Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol Ecol* 28(21):4737–4754
- Stronen AV, Iacolina L, Ruiz-Gonzalez A (2019) Rewilding and conservation genomics: how developments in (re)colonization ecology and genomics can offer mutual benefits for understanding contemporary evolution. *Glob Ecol Conserv* 17:e00502
- Svenning J-C, Pedersen PBM, Donlan CJ, Ejrnæs R, Faurby S, Galretti M, Hansen DM, Sandel B, Sandom CJ, Terborgh JW, Vera FWM (2016) Science for a wilder Anthropocene: synthesis and future directions for trophic rewilding research. *Proc Natl Acad Sci USA* 113(4):898–906
- Turtle Taxonomy Working Group, Rhodin A, Iverson GJ, Bour R, Fritz U, Georges A, Shaffer HB, van Dijk PP (2017) Turtles of the world: annotated checklist and Atlas of taxonomy, synonymy, distribution, and conservation status, (8th edn). In: Rhodin AGJ, Iverson JB, van Dijk PP, Saumure RA, Buhlmann KA, Pritchard PCH, Mittermeier RA (eds) Conservation biology of freshwater turtles and tortoises: a compilation project of the IUCN/SSC tortoise and freshwater turtle specialist group, vol 7. Chelonian Research Monographs, pp 1–292. <https://doi.org/10.3854/crm.7.checklist.atlas.v8.2017>
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA (2013) From fastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinform* 43(1):11.10.1–11.10.33

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.